EDITORIAL

The Relationship between Dynamic Programming and Pontryagin's Minimum principle in Optimal Control

Abdelrahim B. Hamid, Associate Professor, Department of Mathematics and Physics, Faculty of Education Hantoub, University of Gezira.

Abstract

Optimal control theory has been used to obtain solutions to a variety of aerospace engineering problems and holds great promise for other problem areas as well. However, much remains to be accomplished. Hopefully, the reader has been stimulated to learn more about optimal control theory and its applications. In this paper, we discussed the relationship between dynamic programming and Pontryagin's minimum principle and the important features of these two techniques.

The minimum principle determines optimal controls that generally lead to a nonlinear two-point boundary deriving the minimum principle from the Hamilton Jacobi-Bellman functional equation. Particular problem solved with applying minimum principle technique and has been discussed. The minimum principle indicated that the only values assumed determined by an optimal control.

Introduction

Applying the minimum principle, or the calculus of variations, to determine optimal controls generally leads to a nonlinear two-point boundary value problem that requires the use of iterative numerical techniques for solution. As noted, these iterative algorithms determine optimal controls in open-loop form. If the state equations of a process are linear (or have been linearized), and the performance measure is a quadratic form, the optimal control *law* can be determined by numerically integrating a matrix differential equation of the Riccati type. An important feature of the variational approach is that the form of optimal controls can be determined; hence, it is necessary only to consider the subset of controls having the appropriate form; this is a significant conceptual and computational advantage.

Dynamic programming is essentially a clever way of examining all of
the candidates for an optimal control law. To do this by direct enumeration of all the possibilities is a horrendous task, but by using the principle of optimality a multiple-stage decision process can be reduced to a sequence of single stage decision processes, and a feasible compu-tational algorithm is obtained. The algorithm consists of solving the functional recurrence equation.

By a direct search among the admissible control values. The presence of state and control constraints generally complicates the application of variational techniques; however, in dynamic programming, state and control constraints reduce the range of values to be searched and there by simplify the solution. Another desirable feature of the dynamic programming approach is that the computational procedure determines the optimal control law. Moreover, since the algorithm makes a direct comparison of the performance measure values associated with all optimal control law candidates, it is ensured that the global, or absolute, optimal control law is obtained. The primary limitation of the dynamic programming approach is the need for large storage capacity in the digital computer when solving problems involving high-order systems.

Although a particular problem may perhaps be solved by applying only
one of the techniques we have discussed, it is often beneficial to use the complementary features of several different approaches. For example, suppose that the minimum principle indicates that the only values assumed by an optimal control are + I, 0, or -1. This knowledge can be used in designating the control values to be tried in obtaining a dynamic programming solution; instead of trying a finite set of controls that satisfy ·$-1 \leq u \leq +1$, we need use only $u = +1, 0$, and -1 as trial control values. As another example, suppose that it is desired to find an optimal control law for a system whose initial state value is known to be in a specified region of the state space. One approach is to determine an optimal trajectory by employing iterative numerical techniques, and then to make use of this trajectory to define a region of the state space in which an optimal control law can be obtained by dynamic programming. By doing this, only a subset of state space values is searched in the dynamic programming solution, and thus the requirements for computer memory and computation time are eased. As a third example, suppose that the variational approach indicates that a singular interval may occur. Nonlinear or dynamic programming may be helpful in determining whether or not singular controls are optimal.

In most applications engineers are required to design a controller, that
is, a device for generating control signals from observations of system outputs. This being the case, three alternatives are

1. An on-line digital computer that calculates optimal control signals as the process evolves, and additional hardware to synthesize the control signals.
2. A special-purpose digital controller to synthesize an optimal control law that has been precomputed off line with a general purpose digital computer.
3. A suboptimal, but easily implemented, controller whose configuration
and parameters have been precalculated with an off- line computer.

Let us consider the implications of each of these alternatives. For many applications it may be difficult to justify economically the presence of an on-line digital computer. In addition, such a controller must necessarily be suboptimal because of the finite time required for computation. In

EDITORIAL

fact, if the system states change too quickly for the computer to keep exemplify the types of problems that are well suited to on-line control computation.

The second alternative has several advantages. All computing is done off-line; hence, the general-purpose computer required is available for solving many problems rather than being devoted exclusively to one system (or a few systems on a time-shared basis). The special purpose digital controller will be much smaller, less expensive, and not as complex as a general purpose digital computer. In addition, since the optimal control law is precomputed, the question of calculations having to keep up with the changing state of the controlled process does not arise. On the debit side, the control computer may require a large amount of storage; however, this need not always be rapid-access storage a small magnetic tape arrangement may be quite acceptable. Notice that in this control scheme the optimal control law must be calculated. Presumably, this is accomplished by using dynamic programming; an alternative approach that relies on linearization of the state cost ate differential equations about a nominal optimal trajectory is discussed. The concept of suboptimal, but easily implemented, controllers is very attractive from a practical point of view. Naturally, the system's performance with a suboptimal controller should be compared with optimal performance;
such a comparison could be the basis for deciding on the acceptability of a proposed suboptimal design. Rejection of a controller design indicates that either the controller configuration needs to be altered, or· that the controller parameters must be adjusted. Efforts to achieve acceptable suboptimal designs have met with limited success so far. The principal difficulty is that a controller may be nearly optimal for some initial conditions, but very poor for others. A suggested method of alleviating this difficulty is to minimize the deviation from optimal response that results when the system assumes its worst possible initial state':!: This leads to aminimax solution of the problem. An alternative approach is to assume a probability distribution for the initial state values and minimize the expected value of the performance measure. Each of these alternatives generally requires that *all* system states be available for generating the control signal; however, it may be necessary to generate control signals using estimates of the state values obtained from noisy observations of system outputs. One method of obtaining (which are optimal in a statistical sense) is to use a Kalman filter state estimates (which are optimal in a statistical sense).

Results

Considered the problem of finding a control u* $\in U$ that causes a system:

$$\dot{x}(t) = a(x(t)\,,u(t)\,,t) \qquad (1)$$

To respond in such a manner that a performance measure of the form:

$$J = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t)dt \qquad (2)$$

Is minimized, in our discussion of dynamic programming we are interest in Hamilton Jacobi-Bellman functional equation:

$$J_t^*(x(t),t) + \min_{u(t)} \left\{ g(x(t),u(t),t) + \left[J_x^*(x(t),u(t),t)\right]^T a(x(t),u(t),t) \right\} = 0 \qquad (3)$$

must be satisfied by an optimal control and its trajectory. *J*(x(t), t)* is the

EDITORIAL

minimum value of the performance measure for a process beginning at time t with an initial state x(t), J_t^* and J_x^* are partial derivatives of J*(x(t), t)with respect to t and x. If (x*(t), t) is a particular point in the state-time space, the H-J-B functional equation tells us that the optimal control value $u^*(t)$, which corresponds to this point, satisfies the relationship:

$$g(x^*(t),u^*(t),t)+\left[J_x^*(x^*(t),t)\right]^T a(x^*(t),u^*(t),t)$$
$$=\min_{u(t)}\left\{g(x^*(t),u(t),t)+\left[J_x^*(x^*(t),t)\right]^T a(x^*(t),u(t),t)\right\} \qquad (4)$$

for all $t\in [t_0, t_f]$. Thus, we can write Eq. (3) for the point (x*(t), t) as

$$J_t^*(x^*(t),t)+g(x^*(t),u^*(t),t)+\left[J_x^*(x^*(t),t)\right]^T a(x^*(t),u^*(t),t)$$
$$=0 \qquad (5)$$

Equation (5) is a first-order partial differential equation. If t_f is fixed and $x(t_f)$ free, the boundary condition is:

$$J^*(x^*(t_f),t_f)=h(x^*(t_f),t_f) \qquad (6)$$

If Pontryagin's minimum principle is applied to the same optimal control problem, we can find that:

$$\dot{x}^*(t)=\frac{\partial \mathrm{H}}{\partial \mathrm{P}}(x^*(t),u^*(t),\mathrm{P}^*(t),t) \qquad (7)$$

$$\dot{\mathrm{P}}^*(t)=-\frac{\partial \mathrm{H}}{\partial \mathrm{P}}(x^*(t),u^*(t),\mathrm{P}^*(t),t) \qquad (8)$$

$$\mathrm{H}(x^*(t),u^*(t),\mathrm{P}^*(t),t)\le \mathrm{H}(x^*(t),u(t),\mathrm{P}^*(t),t) \qquad (9)$$

for all admissible $u(t)$, and for all $t\in [t_0, t_f]$, are necessary conditions for u* to be an optimal control and x* an optimal trajectory. The boundary conditions for the 2n first-order state-costate differential equations (7) and (8) are:

$$x^*(t_0)=t_0 \qquad (10)$$

$$and\ \ \mathrm{P}^*(t_f)=\frac{\partial h}{\partial x}(x^*(t_f),t_f) \qquad (11)$$

Using the definition of the Hamiltonian:

$$\mathrm{H}(x(t),u(t),\mathrm{P}(t),t)\underline{\underline{\Delta}}\, g(x(t),u(t),t)+\mathrm{P}^T(t)\left[a(x(t),u(t),t)\right],$$

we can write Eqs. (7) and (8) as:

$$\dot{x}^*(t)=a(x^*(t),u^*(t),t) \qquad (12)$$

$$\dot{\mathrm{P}}^*(t)=-\left[\frac{\partial a}{\partial x}(x^*(t),u^*(t),t)\right]^T \mathrm{P}^*(t)-\left[\frac{\partial g}{\partial x}(x^*(t),u^*(t),t)\right] \qquad (13)$$

From the definition: $\Psi_i(x^*(t),t)\underline{\underline{\Delta}}\, J_{x_i}^*(x^*(t),t)$, $i=1,2,3,\cdots,n$ and some properties, we can summarize and show that:

EDITORIAL

$\frac{d\Psi}{dt}(x^*(t),t) = -\left[\frac{\partial a}{\partial x}(x^*(t),u^*(t),t)\right]^T \Psi(x*(t),t) - \frac{\partial g}{\partial x}(x^*(t),u^*(t),t)$ (14) where the optimal control u*(t) satisfies:

$$g(x^*(t),u^*(t),t) + \left[\Psi(x^*(t),t)\right]^T a(x^*(t),u^*(t),t)$$

$$= \min_{u(t)}\left\{g(x^*(t),u(t),t) + \left[\Psi(x^*(t),t)\right]^T a(x^*(t),u(t),t)\right\}$$

and the boundary condition: $\Psi(x^*(t_f),t_f) = \frac{\partial h}{\partial x}(x^*(t_f),t_f)$.

Conclusion

We have succeeded in deriving the minimum principle from the Hamilton Jacobi-Bellman functional equation; however, it is important to keep in mind the restrictions imposed by the derivation. In writing Eqs. (1-14), we assumed that the states are not constrained by any boundaries. Also in provided a useful interpretation of the costate variables.

The primary limitation of the dynamic programming approach is the need for large storage capacity in the digital computer when solving problems involving high-order system.

References

[1] Kelley, H. J., "Method of Gradients," *Optimization Techniques with Applications To Aerospace Systems,* G. Leitmann, ed. New York: Academic Press Inc., 1992.

[2] Kelley, H. J., "Gradient Theory of Optimal Flight Paths," *ARS Journal* (1980)

[3] Lapidus, L., and R. Luus, "The Control of Nonlinear Systems; Part II: Convergence by Combined First and Second Variations," *A. I. Ch. E. Journal*(1997)

[4] Rosen, J. B., "Iterative Solution of Nonlinear Optimal Control Problems,"

EDITORIAL

J. SIAM Control Series A (2004)
[5] Rosen, J. B., "Optimal Control and Convex Programming," *Nonlinear Programming*, J. Abadie, ed. New York: John Wiley & Sons, Inc., 1967.
[6] Sage, A. P., *Optimum Systems Control*. Englewood Cliffs, N. J.: Prentice-Hall, Inc., (2013).